



Ecological niche modelling of an invasive alien plant and its potential biological control agents

P.D. Trethowan^a, M.P. Robertson^{b,*}, A.J. McConnachie^c

^a Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

^b Centre for Invasion Biology, Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

^c Plant Protection Research Institute, Agricultural Research Council, Private Bag X6006, Hilton 3245, South Africa

Received 6 December 2009; received in revised form 9 July 2010; accepted 15 July 2010

Abstract

Invasive alien plants are of concern in South Africa. Pompom weed (*Campuloclinium macrocephalum*) is currently invading the Grassland and Savannah biomes of South Africa and is likely to continue spreading in the southern African sub-region. Two possible biological control agents (*Liothrips tractabilis* and *Cochylys campuloclinium*) have been identified for control of pompom weed. We used ecological niche modelling to predict which areas in southern Africa are likely to be suitable for pompom weed and the two potential biological control agents. The overlap between areas predicted to be highly suitable for pompom weed and areas suitable for the biological control agents was assessed. Methods of reducing sampling bias in a data set used for calibrating models were also compared. Finally, the performance of models calibrated using only native range data, only invaded range data and both were also compared. Models indicate that pompom weed is likely to spread across a greater region of southern Africa than it currently occupies, with the Savannah and Grassland biomes being at greatest risk of invasion. Poor overlap was found between the areas predicted to be highly suitable for pompom weed and those areas predicted to be suitable for the biological control agents. However, models of the potential distribution of the biological control agents are interpreted with caution due to the very small sample size of the data set used to calibrate the models. Models calibrated using both native range and invaded range data were found to perform best whilst models calibrated using only native range data performed the worst. There was little difference found between models that were calibrated using spatially reduced (selecting only one record per 30 min grid cell) and randomly reduced (randomly selecting 50% of available records) biased data sets. © 2010 SAAB. Published by Elsevier B.V. All rights reserved.

Keywords: Biological control; Biological invasions; *Campuloclinium macrocephalum*; Ecological niche modeling; Pompom weed

1. Introduction

Invasive species have resulted in a worldwide conservation and ecological crisis as they enter and alter communities across the globe (Gurevitch and Padilla, 2004). In the USA Pimentel et al. (2005) estimated the economic damage of the effects of alien invasive species to be \$120 billion/year. South Africa seems to be particularly vulnerable to invasive plant species, having one of the biggest problems with alien plant invaders in the world (Richardson and Van Wilgen, 2004).

Pompom weed (*Campuloclinium macrocephalum*) is a perennial herb native to South America and is a member of the family Asteraceae (Henderson et al., 2006). Pompom weed is currently invading the grassland biome of South Africa (Henderson et al., 2006), which is already the biome with the most invasive plant species in South Africa (Richardson and Van Wilgen, 2004). Some of the most threatened vegetation types in South Africa are found in the grassland biome and an estimated 60 to 80% of the biome has already been irreversibly transformed (Knobel and Bredenkamp, 2006). Pompom weed is thought to first invade disturbed areas before spreading to natural grasslands where it is likely to displace native species, resulting in reduced biodiversity and carrying capacity (Henderson et al., 2006).

* Corresponding author.

E-mail address: mrobertson@zoology.up.ac.za (M.P. Robertson).

Several aspects of the ecology of pompom weed contribute to its invasive nature. It has a fleshy, perennial root system that allows it to survive winter frost and fires by storing energy and nutrients underground and re-growing in spring (Henderson et al., 2006). In addition to this, it is drought tolerant, capable of rapid growth and it produces prolific amounts of seed which are effectively dispersed by wind (Henderson et al., 2006), vehicles and movement of cut grass from infested road verges (McConnachie, pers. obs.).

The problem of invasive alien plants that are likely to have an adverse economic or environmental effect on their invaded ranges are frequently addressed by the careful and deliberate introduction of host-specific insects or plant pathogens that reduce the invasive plant's ability to invade (Malecki et al., 1993; McEvoy and Coombs, 1999; McFayden, 1998; Van Wilgen et al., 2004). This method of control is known as biological control. Thus far two potential biological control agents have been identified for pompom weed, namely a flower-feeding lepidopteran, *Cochylis campuloclinium* (Tortricidae) and a stem-galling thrips, *Liothrips tractabilis* (Thripidae). Both of these agents are of South American origin.

Ecological niche modelling can be used to predict the potential geographic distribution of an organism (De Meyer et al., 2008; Guisan and Thuiller, 2005; Mau-Crimmins et al., 2006). This is done by quantifying the species–environment relationship by determining the correlation between the occurrence of the species and the environmental parameters in those regions (Guisan and Thuiller, 2005). In this study we attempt to use ecological niche modelling to predict the potential distribution of pompom weed and its two potential biological control agents. The mapping of the distribution of alien invaders is a popular application of ecological niche modelling (Peterson et al., 2003a,b; Steiner et al., 2008; Zenni et al., 2009). This may allow for management actions to be taken in areas that are predicted to be suitable for the invasive organism (Mau-Crimmins et al., 2006; Mgidi et al., 2007; Rouget et al., 2004; Steiner et al., 2008; Thuiller et al., 2005; Zenni et al., 2009).

Models may be calibrated with native range data, invaded range data or a combination of both. Models calibrated using data from both the native range of the species and the invaded range can provide a better indication of a species potential to spread (Mau-Crimmins et al., 2006). Some attempts to model the range of a species may be restricted to only native range or only invaded range data sets (Baret et al., 2006; Dullinger et al., 2009; Zenni et al., 2009). For example Zenni et al. (2009) modelled the potential distribution of *Acacia paradoxa* in South Africa, however, at this point in time *A. paradoxa* only occurs in a single area in Table Mountain National Park and so only native range data from Australia could be used to calibrate the model. In contrast, Dullinger et al. (2009) restricted the data used in model calibration to invaded range data so as to only capture the niche of the naturalised population.

Models can only be as good as the data that are used to calibrate them (Fielding and Bell, 1997; Hirzel et al., 2001; Wolmarans et al., 2010). This often constitutes a problem when predicting the potential range of invasive species. Acquisition of native range data can be slow, expensive and the researcher may face difficulties, such as literature in foreign languages (Mau-Crimmins et al., 2006).

Researchers attempting to collect data themselves may face time and budget constraints, forcing them to over sample in certain areas whilst ignoring others (and may be restricted by international boundaries). Data collected from various other sources is also subject to a variety of biases depending on why the distribution records were originally collected and who collected them (Funk and Richardson, 2002; Hortal et al., 2008; Reddy and Dávalos, 2003; Wolmarans et al., 2010). In the invaded range, most records tend to be collected where high densities of people occur and in 2003 Reddy and Dávalos showed that historically the intensity of sampling is strongly influenced by accessibility. This means that many data sets are biased samples because the methods used to collect the distribution points did not allow for an equal probability of each point being collected. As a result the dataset does not represent a random sample of the overall distribution of the species that is being sampled. Sampling bias can be detrimental to ecological niche models because the data used to calibrate the model does not represent the climatic niche of the species (Wolmarans et al., 2010). This is true to a large extent for the data used in this study especially the data collected in the invaded range.

The aims of this study were: (1) to predict the potential distributions of pompom weed and the two potential biological control agents in southern Africa using ecological niche models; (2) to assess the overlap in regions predicted to be highly suitable for pompom weed and the two agents; (3) to compare the performance of models for pompom weed that were calibrated with native records only, invaded range records only and a combination of both and (4) to compare two methods of reducing sampling bias in the dataset of occurrence records obtained from the invaded range of pompom weed.

2. Material and methods

2.1. Occurrence records

As many occurrence records as possible were sourced from South America for all three target species. Records for the two biological control agents were obtained during research field trips to South America. The records for pompom weed were obtained from several sources. These included online databases including the Global Biodiversity Information Facility (GBIF, <http://www.gbif.org/>), Tropicos (<http://www.tropicos.org/>) and The New York Botanical Garden (<http://sciweb.nybg.org/science2/vii2.asp>). Other records were obtained from the literature (Biondi et al., 2007; Cavararo, 2004; Cervi et al., 2007; Linsingen et al., 2006; Reginato et al., 2008), from collections data (Museo de La Plata) and during research field trips to South America. Records for pompom weed from its invaded range in southern Africa were obtained from the Southern African Plant Invaders Atlas database (Henderson, 2007).

Maps of occurrence data for the three species were produced to check for obvious errors. To avoid pseudo-replication only one occurrence record per 10 min grid cell was used for model calibration. A total of 187 records were available for pompom weed from its native range (South America) and 130 records

from its invaded range (southern Africa). As the biological control agents have not been released in South Africa, records were available only from their native ranges (South America): 26 records for *C. campuloclinium* and 16 records for *L. tractabilis*.

2.2. Predictor variables

Suitable bioclimatic predictor variables were obtained from the Worldclim database (Hijmans et al., 2005). The same predictor variables as used by Beaumont et al. (2009) were used as predictor variables in this study. According to Beaumont et al. (2009) these include an uncorrelated range of averages, outliers and seasonal variation. The predictor variables included mean diurnal temperature range, maximum temperature of the warmest month, mean temperature of the coldest quarter, annual precipitation, precipitation of the driest month and precipitation seasonality. A suitable spatial resolution was chosen (10 min) and the data for the predictor variables was obtained at this resolution.

2.3. Ecological niche modelling

Ecological niche models for all three species were developed using Maxent (Phillips and Dudík, 2008; Phillips et al., 2006). Maxent is a popular correlative modelling technique that has performed favourably when compared with other techniques (Elith et al., 2006). The advantage of using Maxent is that it can make use of presence-only data, as opposed to presence–absence data. It has been used successfully for ecological niche modelling of invasive species (e.g. Steiner et al., 2008; Wolmarans et al., 2010).

2.4. Model calibration and evaluation

2.4.1. Biological control agent models

Models were calibrated using all the native range records for each species to predict their potential invaded ranges in southern Africa. This approach was followed so that the maximum number of records could be used in model calibration for models projected to southern Africa to ensure that potential ranges were as accurate as possible. However, this meant that these models could not be evaluated as these species have not been released in southern Africa. As a result we randomly split the full dataset into a calibration set (70%) for model calibration and an evaluation set (30%) that was used to evaluate model performance. We performed twenty random splits by randomly sampling the 70%:30% subset with replacement. We calculated the Area Under Curve (AUC) statistic for each replicate model to evaluate model performance.

The AUC statistic is derived from the Receiver Operating Characteristic (ROC) plot, which is a measure of performance that provides a single threshold-independent measure of accuracy (Fielding and Bell, 1997). Usually the calculation of the AUC statistic requires both presence and absence data, but in many cases absence data are not available. As a result the AUC statistic is calculated in Maxent by using presence and background data (Phillips and Dudík, 2008; Phillips et al.,

2006). The values of the AUC range from 0.5 to 1.0, with 0.5 being no better than random (Fielding and Bell, 1997). Although there are no strict rules for classification Beaumont et al. (2009) considered an AUC of less than 0.8 to be a poor model, an AUC between 0.8 and 0.9 a fair model, between 0.9 and 0.95 was considered to be a good model and greater than 0.95 to be an excellent model.

2.4.2. Pompom weed models

For pompom weed, models were calibrated using different datasets which yielded five different treatments. Models were calibrated using: all native range records (N), invaded range records (I), native and invaded range records (NI), all native range records and a random sample of 50% of the invaded range records (NIr), and with all native range records and by selecting one record per 30 min grid cell from the invaded range (NIs). The NIr and NIs treatments were undertaken in order to try and reduce the effects of sampling bias in the invaded range data set. The original set of invaded range records was randomly split into a calibration set (70%) for model calibration and an evaluation set (30%) that was used to evaluate model performance. Background data that are required by Maxent for model calibration (Phillips and Dudík, 2008; Phillips et al., 2006) were taken from South America only. The models were then projected to southern Africa. Twenty replicates were performed per treatment by randomly sampling the 70%:30% subset with replacement so that different records could be used in testing and training.

We calculated the AUC statistic for each replicate model to evaluate model performance. Average AUC statistics were calculated for each treatment to assess model performance. Analysis of variance (ANOVA) was used to compare AUC values among treatments. A Tukey HSD test was then performed to determine which treatments differed significantly from one another.

2.5. Invasion risk

We produced a simple model of pompom weed invasion risk. The model consisted of three components: 1) climate suitability of pompom weed, 2) land cover transformation, and 3) distance from nearest pompom occurrence record. We took the map from the best performing NI model to represent the climate suitability. We assigned all transformed and degraded land classes in the National Land Cover classification of South Africa, Lesotho and Swaziland of 2000 to a value of 1. All natural land cover classes were assigned a value of 0. This resulted in the land cover transformation map. We calculated the Euclidean distance to the nearest pompom weed occurrence record (using ArcGIS Euclidean distance module) for each grid cell in a raster that had the same spatial resolution as the land cover map. We then rescaled the Euclidean distances between 0 and 1 so that the furthest distance in the map had a value of 0 and the closest distance to a pompom weed occurrence had a value of 1. This we call the distance map. Finally, we produced an invasion risk map by adding the climate suitability map, the land cover transformation map and the distance map; and then

rescaling the values between 0 and 1. The index of invasion risk ranged between 0 (low risk) and 1 (high risk). All calculations were performed in ArcGIS v9.3.

2.6. *Overlap of potential distributions*

The overlap between the potential distribution of the biological control agents and the predicted potential distribution of pompom weed was assessed by comparing their potential distribution maps. To do this we converted the best performing NI model for pompom weed to a presence–absence map using the probability value associated with the 10th percentile of training presence (which is calculated by Maxent). This threshold was selected as it gave a good overlap with known occurrences. We also converted the potential distribution maps of the two biological control agents into presence–absence maps using a similar approach. For these we used the fixed cumulative value 5 logistic threshold calculated by Maxent for each model because the 10th percentile of training presence threshold gave what appeared to be very conservative maps. We superimposed the biological control agent maps with the pompom weed map to determine areas of overlap. We calculated the proportional overlap (Reyers et al., 2000) between the biological control agent maps and the pompom weed map. Proportional overlap was calculated by taking the number of grid cells of overlap and dividing by the maximum number of overlapping grid cells possible (i.e. the map with the smallest number of grid cells).

3. Results

3.1. *Pompom weed*

In southern Africa pompom weed has been recorded most frequently in the Grassland and Savannah biomes (Fig. 1). Most records are from the highveld grasslands in the Gauteng Province, as sampling effort seems to have been highest in this

area. The plant has also been recorded in Mpumalanga, Limpopo, Northwest and Free State provinces (Fig. 1). A confirmed occurrence was also reported near George in the Western Cape (Fig. 1a). Fig. 1a shows all occurrence records of pompom weed, but to avoid pseudo-replication we only used one record per 10 min grid cell to calibrate the models. This reduced the number of records used for model calibration. The number of records used in each of the treatments is summarised in Table 1. Once the occurrence records had been randomly reduced by 50% (N_Ir treatment), 45 occurrence records remained in the invaded range for pompom weed. After spatially reducing the occurrence records using a 30 min grid (N_Is treatment), 37 occurrence records remained in the invaded range of pompom weed. Far fewer invaded range occurrence records were used in the N_Ir and N_Is treatments than were used in the NI and I treatments.

Regions of highest suitability for pompom weed included the Grassland and Savannah Biomes in the Gauteng, Mpumalanga, Limpopo, Free State, KwaZulu-Natal and North West Provinces of South Africa and also parts of Lesotho, Swaziland, Mozambique and Zimbabwe (Figs. 1–3). The potential distribution maps varied according to the data that were used to calibrate the models (Figs. 2 and 3). The model calibrated using data from the invaded range only (I, Fig. 2a) showed lower probabilities over the map region than models calibrated with native range data only (N, Fig. 2b) or a combination of native and invaded range data (NI, Fig. 2c). The regions of highest probability coincided quite well with those regions where the plant has been recorded (Fig. 1a) for the I and NI models (Figs. 2b, c) but not for the N models (Fig. 2a). The NI models (Fig. 2c) predicted a much larger region of high probability (>0.4) than the I models (Fig. 2b).

The models calibrated using native range data with randomly reduced invaded range data (N_Ir, Fig. 3a) and spatially reduced data (N_Is, Fig. 3b) produced potential distributions that were similar to the potential distribution for models calibrated with native and invaded data that was not reduced (NI, Fig. 2c).

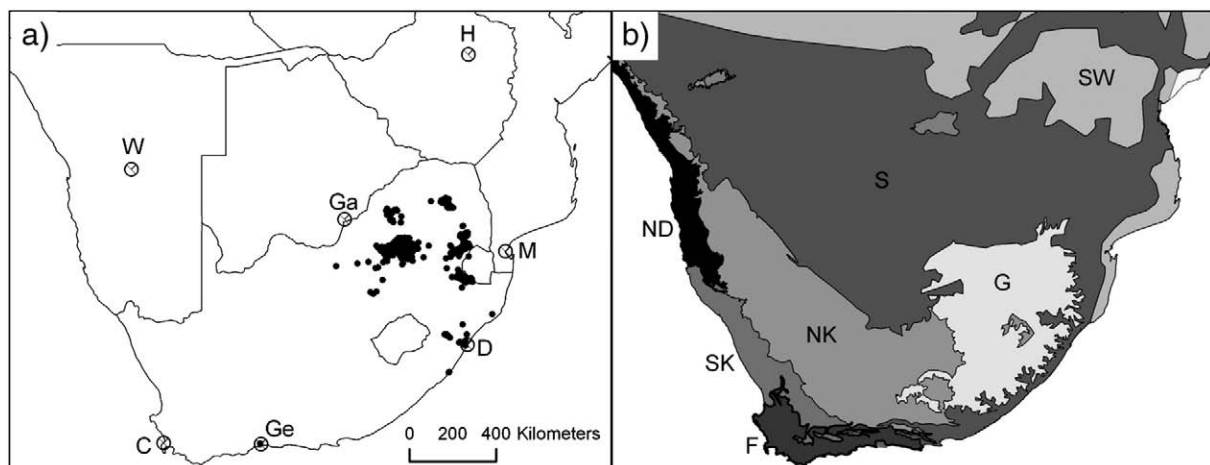


Fig. 1. The invaded range of pompom weed (*Campuloclinium macrocephalum*) in southern Africa, a) occurrence records and cities: C – Cape Town, Ge – George, D – Durban, M – Maputo, Ga – Gaborone, W – Windhoek, H – Harare; b) The biomes of southern Africa (White 1983): F – Fynbos, SK – Succulent Karoo, NK – Nama Karoo, ND – Namib Desert, G – Grasslands, S – Tree and shrub savanna, SW – subtropical woodland.

Table 1

The number of occurrence records used to calibrate the pompom weed models for each of the treatments.

Description	Treatment	Native	Invaded	Total
Native records only	N	187	0	187
Invaded records only	I	0	91	91
Native and invaded	NI	187	91	278
Native and invaded, randomly reduced	NIr	187	45	232
Native and invaded, spatially reduced	NIs	187	37	224

There was a significant difference in model performance across treatments based on their mean AUC statistic ($F_4=1461.7$ $p<0.05$). The subsequent Tukey HSD test demonstrated that in terms of the AUC statistic, models calibrated with only native range data (N) performed significantly worse (mean AUC of 0.764) than all other models (I, NI, NIr, NIs, Fig. 4). This is supported by the potential distribution maps that showed quite low probabilities (0.2) for N (Fig. 2a) for most of the region where the species has been recorded (Fig. 1a). In contrast, the potential distributions for the other models (treatments: I, NI, NIr and NIs) had high probabilities in regions where the species has been recorded (Figs. 2 and 3).

Models calibrated using either invaded range data only (I) or both invaded and native range data (NI, means of 0.963 ± 0.01 and 0.966 ± 0.013 respectively) did not differ significantly in performance but do perform significantly better than the NIr and NIs models (Fig. 4). Models calibrated using native range data together with invaded range data that has been reduced either spatially or randomly (means of 0.943 ± 0.012 and 0.936 ± 0.012 respectively) did not perform significantly differently from each other. However, these models (NIr and NIs) performed significantly better than models using only native range data (N), but worse than the NI models (Fig. 4).

The invasion risk model identified several areas in South Africa that can be considered to have a high risk of invasion by pompom weed (Fig. 5).

3.2. Biocontrol agents and range overlap

The average AUC value for the 20 replicate models for *C. campuloclinium* was 0.980 and 0.973 for *L. tractabilis* (evaluated using 30% of native range data). Model performance can be considered to be excellent for both species. The potential distribution of *C. campuloclinium* is concentrated in the arid Northern Cape of South Africa, southern Namibia and southern Botswana (Fig. 6b), mostly the Nama Karoo (Fig. 1b). The potential distribution of *L. tractabilis* is centred on the same region as that of *C. campuloclinium* but occurs more widely in Eastern and Western Cape provinces of South Africa (Fig. 6c). There was very poor overlap between the potential distributions of the biological control agents and the potential distribution of pompom weed (Fig. 6b, c). Proportional overlap between the *C. campuloclinium* map and the pompom weed map (Fig. 6b) was very low at only 0.001. The proportional overlap between the *L. tractabilis* map and the pompom weed map (Fig. 6c) was higher at 0.018.

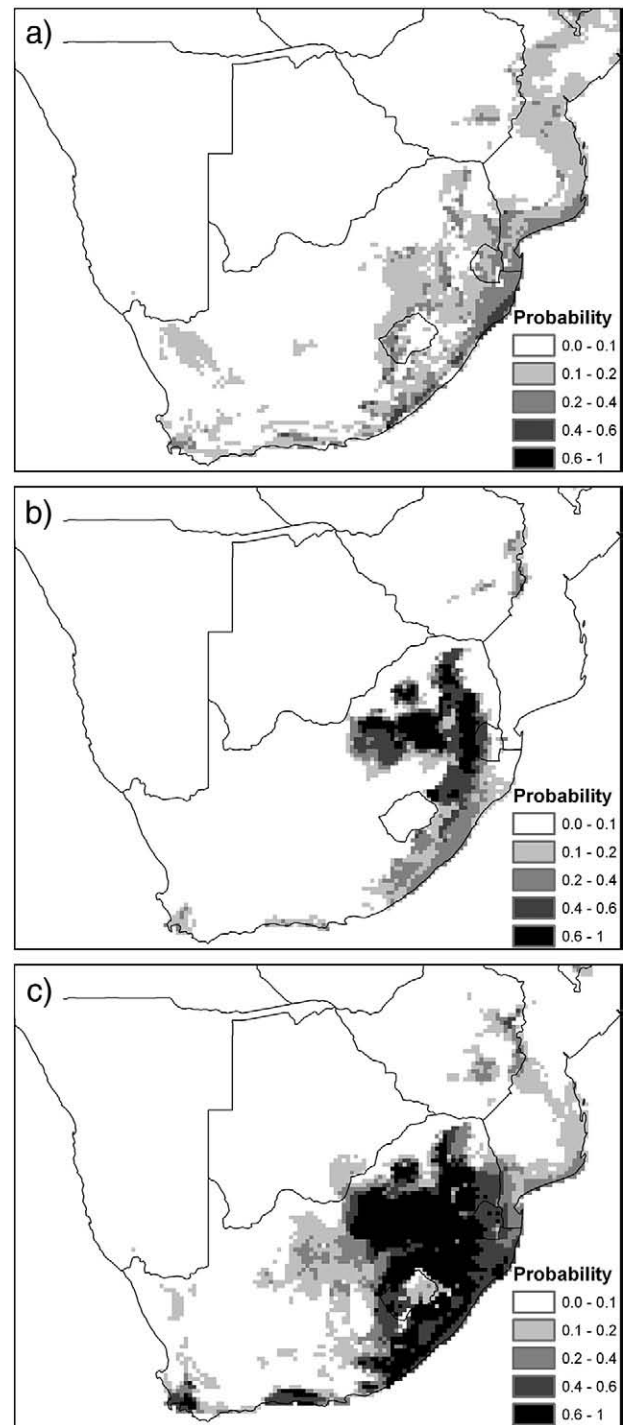


Fig. 2. The potential distribution of pompom weed in southern Africa calibrated using (a) native, (b) invaded, and (c) native and invaded range occurrence records. Maps show the average of 20 replicate models.

4. Discussion

4.1. Pompom weed potential distribution and invasion risk

It appears that in the absence of appropriate interventions, pompom weed could spread from its current distribution across a large portion of southern Africa. The potential distribution maps indicate the potential distribution of pompom weed based

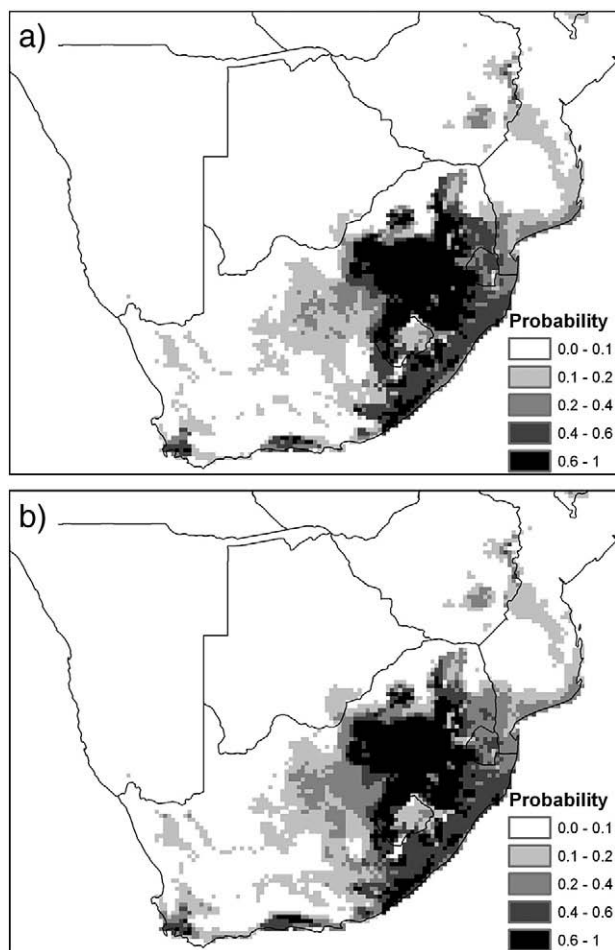


Fig. 3. The potential distribution of pompom weed (a) calibrated using native range data and randomly reduced invaded range data and (b) calibrated using native range data and spatially reduced invaded range data. Maps are means of 20 replicate models.

on climate suitability. In order to understand invasion risk, additional factors need to be considered (Richardson and Van Wilgen, 2004). Our simple model of pompom weed invasion risk incorporates potential suitability of climate for pompom weed, distance to nearest pompom occurrence record and

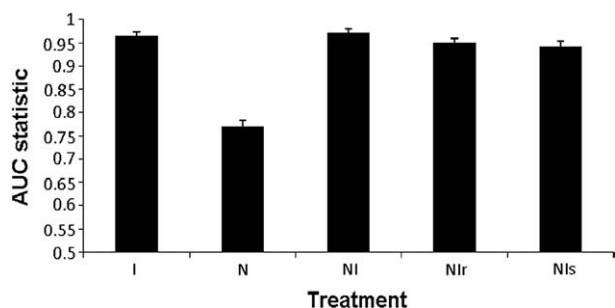


Fig. 4. The mean Area Under Curve (AUC) statistic and standard deviations from 20 replicate models of pompom weed for five treatments; I – models calibrated using only invaded range data, N – only native range data, NI – native range and invaded range data, NIr – native range and randomly reduced invaded range data, NIs – native range and spatially reduced invaded range data. Treatments with different letters differed significantly ($p < 0.05$).

transformed land cover (disturbance). Invasion risk is likely to be highest in areas that are climatically most suitable, that are disturbed and that are in close proximity to sources of propagules (Richardson and Van Wilgen, 2004). Our simple risk model did not consider disturbance associated with roadsides, as the available land cover map was not produced at this fine resolution. Roadsides are likely to be important for invasion of this species as they represent linear disturbances. This could have resulted in an underestimation of invasion risk.

Based on the risk model, many areas in South Africa appear to be at risk from invasion by pompom weed. Some of the areas identified as being climatically suitable for pompom weed that are also transformed, are small and are separated by large distances from invaded areas (e.g. patches of high suitability located in the Eastern Cape, the Western Cape). Although these areas are considered highly suitable climatically for pompom weed, they are not necessarily at high risk of invasion in the immediate future.

It is also clear that there are many areas that are at risk of invasion that have not been invaded. This provides a good opportunity to put in place pre-emptive measures to prevent invasion. However, pompom weed produces many wind-dispersed seeds that appear to be easily dispersed and this may hamper any efforts to prevent invasion.

4.2. Model calibration

Mau-Crimmins et al. (2006) and Beaumont et al. (2009) evaluated the importance of calibrating models using either native range data or invaded range data. Models tend to perform better when calibrated using invaded range data than when calibrated using native range data (Beaumont et al., 2009; Mau-Crimmins et al., 2006). Models calibrated using invaded range data can consider factors not present in the native range, such as climatic preferences of the invaded range genotype (Mau-Crimmins et al., 2006). The results here support the findings of Mau-Crimmins et al. (2006) and Beaumont et al. (2009) that models calibrated using invaded range data perform better than models calibrated using native range data. The models calibrated using native range data had a mean AUC statistic of less than 0.8. This is classed as a poorly performing model according to Beaumont et al. (2009). In contrast, the model calibrated using at least some invaded range data (I, NI, NIr and NIs treatments) all had mean AUC statistics above 0.95 (classed as excellent model performance by the Beaumont et al., 2009) except for the NIs treatment which had a mean AUC statistic of greater than 0.9 (classed as a model that performs well according to Beaumont et al., 2009).

It appears that the native range model did not capture the ecological niche of pompom weed effectively as most known presences had quite low probability values. The best performing models are those that were calibrated using both native range data and invaded range data. This demonstrates the importance of obtaining both data sets when attempting to map the potential distribution of an invasive plant (Beaumont et al., 2009).

There were many occurrence records for pompom weed from the invaded range (813 records) that were largely focussed

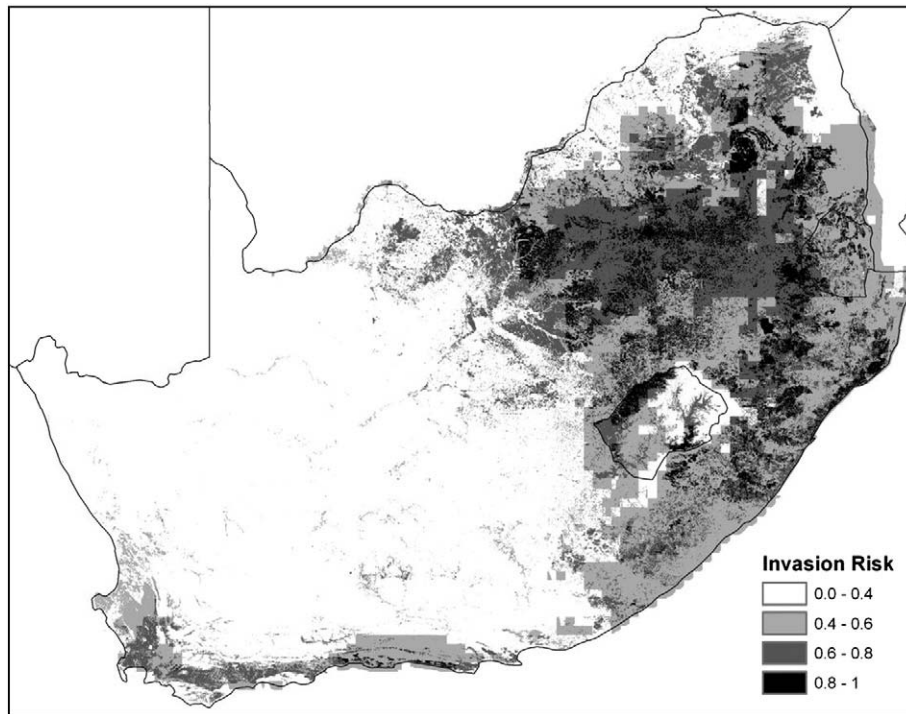


Fig. 5. Map of invasion risk for pompom weed in South Africa, Lesotho and Swaziland.

in a relatively small area (primarily restricted to two or three provinces in South Africa) whereas there were relatively few native range occurrence records for pompom weed (288) scattered over much of the continent of South America. For this reason we believed that the invaded range was oversampled (in that a disproportionately large number of records had been collected from a small area) and we compared methods of reducing the invaded range sampling bias. To do this we compared two methods of reducing sampling bias in the invaded range. In the first method we randomly reduced the data set by a certain percentage (50%). In the second we used a spatial method of reducing sampling bias. This was done by only choosing one record per 30 min grid cell. We expected that spatially reducing a biased data set would be a better method of reducing sampling bias than randomly reducing the data set. When randomly reducing a data set, it is highly likely that areas containing only a few peripheral records will be removed from the data set. Peripheral records may represent a novel climatic region that may be important to the model and affect its outcome. Spatially reducing the data set will not remove these records. By spatially reducing a data set, sampling bias can therefore be reduced whilst still covering the full spectrum of climatic regions occupied by the species. However, we did not find that either method of reducing the training data or not reducing the training data at all (beyond the original 10 min grid) made much difference to the models. There was no significant difference in mean AUC statistics for these models and they predicted similar areas as being suitable for pompom weed. Although in this case model performance decreased when the invaded range dataset was reduced (either spatially or randomly) it may still be sensible in other circumstances to

reduce a biased dataset. The reduction of the invaded range data set here may have had a greater effect if we had used finer resolution predictor variables (e.g. 5 min) because fewer records would have been excluded in the initial filtering of the records than were removed by the 10 min grid.

The native range model may not have captured the ecological niche of pompom weed due to two main reasons; it is possible that niche shift has occurred, or that our native range data did not capture the full range occupied by pompom weed in South America. To understand niche shift it is important to distinguish between the fundamental and realised niche. The fundamental niche of a species is the region in which the abiotic factors are suitable for the species, and the realised niche is usually a smaller subset of this, in which both the abiotic and the biotic factors are suitable enough in order to allow the species to occur there (Hutchinson, 1957; Soberón and Peterson, 2005). Ecological niche models that are calibrated using occurrence records are thus based on the realised niche rather than the fundamental niche. Niche shift is defined as a change in position of either the fundamental or realised niche of a species (Pearman et al., 2008). A shift in the realised niche can occur due to the release from competitors and natural enemies in the invaded range. If genetic changes to environmental tolerances of the species occur in the invaded range of the species then this can result in changes in both the realised and fundamental niches (Pearman et al., 2008).

Niche shift has been reported in certain organisms when moving from their native range to a new invaded range (Broennimann et al., 2007; Fitzpatrick et al., 2007). Ecological niche models calibrated using only native range data rely on the assumption that a species conserves its ecological niche when invading a new range (Broennimann et al., 2007; Pearman et al.,

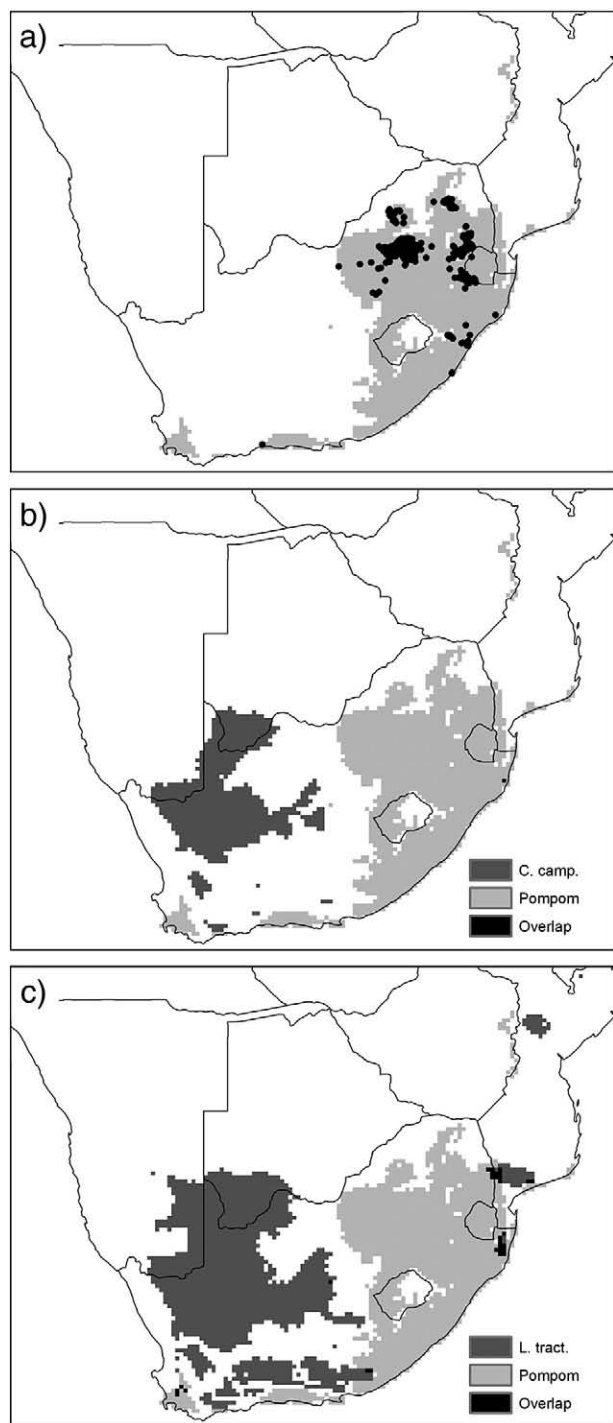


Fig. 6. The potential distribution of (a) the best performing model of pompom weed (AUC=0.984) transformed to a presence–absence map using the 10th percentile of training presence as the threshold value (0.325); (b) The potential distribution of *Cochylis campuloclinium* transformed to a presence–absence map using the fixed cumulative value 5 logistic threshold value (0.088) and (c) *Liothrips tractabilis* transformed to a presence–absence map using the fixed cumulative value 5 logistic threshold value (0.070).

2008). However, Broennimann et al. (2007) showed that certain plant species can occupy ecologically distinct niches in their native and invaded ranges. This may be the case with pompom weed and is possibly the cause of the poor performance of the models calibrated using only native range data.

The second possible reason for poor performance of the models calibrated with native range data only is that the native range data may not have captured the full extent of the native range of pompom weed (e.g. if parts of the species' range were not sampled). Therefore the realised niche of pompom weed could be under represented, resulting in models calibrated with this data to omit climatically suitable areas from the prediction of the potential range of the species.

The best performing pompom weed models were those calibrated using both native and invaded range data. This supports the idea that the native range data is important for the model even though it may not necessarily capture the full extent of the climatic conditions that are suitable for pompom weed.

4.3. Biocontrol agents and range overlap

There is a very poor overlap between the potential distributions of the biological control agents the potential distribution of pompom weed. This suggests that these potential biological control agents may not be well suited to control pompom weed in southern Africa. However, the poor overlap is more likely to do with the quality of the datasets available for calibrating the models for the biological control agents. The sample sizes of these datasets are small and it is likely that these occurrence records underestimate the native ranges of these species. Better native range datasets are required in order to improve our ability to model the potential distribution of these biological control agents accurately. There is likely to be a much better overlap between the biological control agent distributions and pompom weed than the models suggest. Thus we cannot reject these agents as being unsuitable for the control of pompom weed due to their climatic preferences until better native range datasets are available. Based on the same climate matching principles used by Robertson et al. (2008) it would probably be best to search in areas that are climatically similar to the areas of southern Africa that are predicted to be highly suitable for pompom weed so that the biological control agents are pre-adapted to local conditions. The potential distribution maps of the two biological control agents are fairly similar to one another. This is not unexpected because the models were calibrated using very similar data sets of occurrence records.

5. Conclusion

It appears that pompom weed will continue to spread to a far greater region of southern Africa than it currently occupies. If this does occur, then suitable biological control agents will need to be released. The solution may lie with *C. campuloclinium* and *L. tractabilis*, however, our distribution data for these species was insufficient to recommend whether or not these species will occupy a climatically similar niche to pompom weed. Preliminary results suggest otherwise and at the very least, it may be useful to use climate matching in order to search for biological control agents in appropriate areas of South America. Although we did not find a significant improvement in model performance when reducing the number of records in the invaded range dataset that was used for model calibration, this

may still be an important factor to consider for other invasive species and when models are produced at finer spatial resolutions. Further, our results are consistent with other studies in the literature that have compared the importance of native range data and invaded range data in finding that invaded range data is more useful than native range data in model calibration. However, when both native and invaded range are available, then both data sets should be used.

Acknowledgments

We thank Lesley Henderson for making data from the Southern African Plant Invaders Atlas available to us. The Working for Water programme (DWEA) is thanked for funding biological control surveys in the native range of *C. macrocephalum*. Fernando McKay (USDA-ARS South American Biological Control Laboratory) is thanked for assistance with logistics and surveys in Argentina. We thank an anonymous reviewer for helpful comments on a previous version of the manuscript.

References

- Baret, S., Rouget, M., Richardson, D.M., Lavergne, C., Ego, B., Dupont, J., Strasberg, D., 2006. Current distribution and potential extent of the most invasive alien plant species on La Réunion (Indian Ocean, Mascarene islands). *Austral Ecology* 31, 747–758.
- Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R., Hughes, L., 2009. Developing climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions* 15, 409–420.
- Biondi, D., Leal, L., Batista, A.C., 2007. Fenologia do florescimento e frutificação de espécies nativas dos Campos. *Acta Scientiarum – Biological Sciences* 29, 269–276.
- Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T., Guisan, A., 2007. Evidence of climatic niche shift during biological invasion. *Ecology Letters* 10, 701–709.
- Cavararo, R., 2004. Ambiente e plantas vasculares. Reserva Ecológica do IBGE.
- Cervi, A.C., Linsingen, L., Hatschbach, G., Ribas, O.S., 2007. A Vegetação do Parque Estadual de Vila Velha, Município de Ponta Grossa, Paraná, Brasil. *Boletim do Museu Botânico Municipal* 69, 1–52.
- De Meyer, M., Robertson, M.P., Peterson, A.T., Mansell, M.W., 2008. Ecological niches and potential geographical distributions of Mediterranean fruit fly (*Ceratitis capitata*) and Natal fruit fly (*Ceratitis rosa*). *Journal of Biogeography* 35, 270–281.
- Dullinger, S., Kleinbauer, I., Peterseil, J., Smolik, M., Essl, F., 2009. Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. *Biological Invasions* 11, 2401–2414.
- Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC, Overton, J., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Sobero'n, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species distributions from occurrence data. *Ecography* 29, 129–151.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38–49.
- Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J., Dunn, R.R., 2007. The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography* 16, 24–33.
- Funk, V.A., Richardson, K.S., 2002. Systematic data in biodiversity studies: use it or lose it. *Systematic Biology* 51, 303–316.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8, 993–1009.
- Gurevitch, J., Padilla, D.K., 2004. Are invasive species a major cause of extinctions? *Trends in ecology and evolution* 19, 470–474.
- Henderson, L., 2007. Invasive, naturalized and causal alien plants in southern Africa: a summary based on the Southern African Plant Invaders Atlas (SAPIA). *Bothalia* 37, 215–248.
- Henderson, L., Goodall, J.M., Klein, H., 2006. Pompom weed (*Campuloclinium macrocephalum*) an invader of grasslands that threatens conservation and agriculture in South Africa. Agricultural Research Council-Plant Research Institute, pamphlet.
- Hijmans, R.J., Cameron, S.E., Parr, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965–1978.
- Hirzel, A.H., Helfer, V., Metral, F., 2001. Assessing habitat suitability models with a virtual species. *Ecological Modelling* 145, 111–121.
- Hortal, J., Jiménez-Valverde, A., Gómez, J.F., Lobo, G.M., Baselga, A., 2008. Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos* 10, 1–12.
- Hutchinson, G.E., 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22, 415–442.
- Knobel, J., Bredenkamp, G., 2006. The magnificent natural heritage of South Africa. Sunbird publishers, Roggerbaai, pp. 74–83.
- Linsingen, L., Sonehara, J.S., Uhlmann, A., Cervi, A.C., 2006. Composição florística do parque estadual do Cerrado de Jaguariaiva, Paraná, Brasil. *Acta Biológica Paranaense* 35, 197–232.
- Malecki, R.A., Blossy, B., Hight, S.D., Schroeder, D., Kok, L.T., Coulson, J.R., 1993. Biological control of Purple Loosestrife. *Bioscience* 43, 680–686.
- Mau-Crimmins, T.M., Schussman, H.R., Geiger, E.L., 2006. Can the invaded range of a species be predicted sufficiently using only native-range data? *Ecological Modelling* 193, 736–746.
- McEvoy, P.B., Coombs, E.M., 1999. Biological control of plant invaders: regional patterns, field experiments, and structured population models. *Ecological Applications* 9, 387–401.
- McFayden, R.E.C., 1998. Biological control of weeds. *Annual review of Entomology* 43, 369–393.
- Mgidi, T.N., Le Maitre, D.C., Schonegevel, L., Nel, J.L., Rouget, M., Richardson, D.M., 2007. Alien plant invasions — incorporating emerging invaders in regional prioritization: a pragmatic approach for southern Africa. *Journal of Environmental Management* 84, 173–187.
- Pearman, P.B., Guisan, A., Broennimann, O., Randin, C.F., 2008. Niche dynamics in space and time. *Trends in Ecology and Evolution* 23, 149–158.
- Peterson, A.T., Papes, M., Kluza, D.A., 2003a. Predicting the potential invasive distributions of four alien plants species in North America. *Weed Science* 51, 863–868.
- Peterson, A.T., Scachetti-Pereira, R., Kluza, D.A., 2003b. Assessment of invasive potential of *Homalodisca coagulata* in western North America and South America. *Biota Neotropica* 3, 1–7.
- Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31, 161–175.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190, 231–259.
- Pimmentel, D., Zuniga, R., Morrison, D., 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52, 273–288.
- Reddy, S., Dávalos, L.M., 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography* 30, 1719–1727.
- Reginato, M., Matos, F.B., Lindoso, G.D.S., Fogaca de Souza, C.M., Prevedello, J.A., Morais, J.W., Evangelesta, P.H.L., 2008. The vegetation in the reserve “Mata Viva”, Curitiba, Paraná, Brazil. *Acta Biologica Paranaense* 37, 229–252.
- Reyers, B., Van Jaarsveld, A.S., Kruger, M., 2000. Complementarity as a biodiversity indicator strategy. *Proceedings of the Royal Society of London Series B* 267, 505–513.

- Richardson, D.M., Van Wilgen, B.W., 2004. Invasive alien plants in South Africa: how well do we understand the ecological impacts? *South African Journal of Science* 100, 45–52.
- Robertson, M.P., Kriticos, D.J., Zachariades, C., 2008. Climate matching techniques to narrow the search for biological control agents. *Biological Control* 46, 442–452.
- Rouget, M., Richardson, D.M., Nel, J.L., Le Maitre, D.C., Egoh, B., Mgid, T., 2004. Mapping the potential ranges of major plant invader in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions* 10, 475–484.
- Steiner, F.M., Schlick-Steiner, B.C., Van Der Waal, J., Reuther, K.D., Christian, E., Stauffer, C., Suarez, A.V., Williams, S.E., Crozier, R.H., 2008. Combined modelling of distribution and niche in invasion biology: a case study of two invasive *Tetramorium* ant species. *Diversity and Distributions* 14, 538–545.
- Soberón, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species distributional areas. *Biodiversity Informatics* 2, 1–10.
- Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O., Rouget, M., 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11, 2234–2250.
- Van Wilgen, B.W., De Wit, M.P., Anderson, H.J., Le Maitre, D.C., Kotze, I.M., Ndala, S., Brown, B., Rapholo, M.B., 2004. Costs and benefits of biological control of invasive alien plants: case studies from South Africa. *South African Journal of Science* 100, 113–122.
- Wolmarans, R., Robertson, M.P., Van Rensburg, B.J., 2010. Predicting invasive alien plant distributions: how geographical bias in occurrence records influences model performance. *Journal of Biogeography*. [10.1111/j.1365-2699.2010.02325.x](https://doi.org/10.1111/j.1365-2699.2010.02325.x)
- White, F.B., 1983. *Ecosystems of the world. Tropical compilation: savannas*. Elsevier Scientific Publishing Company, Amsterdam. No.13.
- Zenni, R.D., Wilson, J.R.U., Le Roux, J.J., Richardson, D.M., 2009. Evaluating the invasiveness of *Acacia paradoxa* in South Africa. *South African Journal of Botany* 75, 485–496.